



# UNITED STATES PATENT AND TRADEMARK OFFICE

68  
UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/493,748	01/28/2000	Kenji C. Obata	MSFT112958	2621

26389 7590 09/25/2003

CHRISTENSEN, O'CONNOR, JOHNSON, KINDNESS, PLLC  
1420 FIFTH AVENUE  
SUITE 2800  
SEATTLE, WA 98101-2347

[REDACTED] EXAMINER

NGUYEN, MAIKHANH

[REDACTED] ART UNIT

[REDACTED] PAPER NUMBER

2176

DATE MAILED: 09/25/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

<b>Office Action Summary</b>	Application No.	Applicant(s)
	09/493,748	OBATA ET AL.
	Examiner	Art Unit
	Maikhahan Nguyen	2176

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

#### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

- 1) Responsive to communication(s) filed on 28 January 2000.
- 2a) This action is FINAL.                  2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

- 4) Claim(s) 1-20 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) Claim(s) \_\_\_\_\_ is/are allowed.
- 6) Claim(s) 1-20 is/are rejected.
- 7) Claim(s) \_\_\_\_\_ is/are objected to.
- 8) Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on \_\_\_\_\_ is/are: a) accepted or b) objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- 11) The proposed drawing correction filed on \_\_\_\_\_ is: a) approved b) disapproved by the Examiner.
- If approved, corrected drawings are required in reply to this Office action.
- 12) The oath or declaration is objected to by the Examiner.

#### Priority under 35 U.S.C. §§ 119 and 120

- 13) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) All b) Some \* c) None of:
1. Certified copies of the priority documents have been received.
  2. Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.
- 14) Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).
- a) The translation of the foreign language provisional application has been received.
- 15) Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

#### Attachment(s)

- |  |   |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)                                  | 4) <input type="checkbox"/> Interview Summary (PTO-413) Paper No(s). _____  |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)                         | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449) Paper No(s) <u>3</u> . | 6) <input type="checkbox"/> Other: _____                                    |

***DETAILED ACTION***

1. This action is responsive to communications: original application filed 01/28/2000; IDS filed 09/04/2002.
2. Claims 1-20 are currently pending in this application. Claims 1 and 10 are independent claims.

***Rejections - 35 USC § 102***

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless -

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2)a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a)shall have the effects for the purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2)of such treaty in the English language; or " (Emphasis added.)

Claims 1-20 are rejected under 35 U.S.C. 102(e) as being anticipated by **Chakrabarti et al.** (U.S. 6,418,433 – filed 01/1999).

**As to independent claim 1,** Chakrabarti teaches (*col.5, lines 14-60*) a computer-implemented method for selectively accessing a document during a current crawl (*a user can search ... Web pages of interest*) of a server computer (*Web server*), the document being identified by a document address specification (*a Web page URL*), the document having been retrieved during a previous crawl (*new page/old page*), the method comprising:

- determining whether to access the document during the current crawl with the aid of a statistical model (*explores the Web starting from the example set, using the statistics collected from the examples and other analysis on the link graph of the growing crawl database; Abstract*); and

- accessing the document if the determination produces an instruction indicative that the document at the document address specification should be accessed during the current crawl (*recognizes Web pages that are relevant to the interest of one or more users ...guide itself towards relevant, valuable resources; Abstract*).

**As to dependent claim 2,** Chakrabarti teaches determining whether to access the document further comprises computing a probability that the document has changed since the document was retrieved during the previous crawl (*evaluate for potential changes to old pages that might have occurred since the last time the old pages were considered by the system; col.8, lines 53-67*).

**As to dependent claim 3,** Chakrabarti teaches selecting an active probability indicative of a proportion of documents in a plurality of documents that are changing at various change rates, the plurality of documents including the document; training the active probability to reflect an experience with the document during a plurality of previous crawls; and using the trained active probability to compute the probability that the document has changed (*col.7, lines 3-27 & col.8, lines 52-67*).

**As to dependent claim 4,** Chakrabarti teaches selecting the probability that the document has changed from the previous crawl as the active probability in the current crawl; and repeating the method of Claim 3 for the current crawl (*col.7, lines 3-27 & col.8, lines 52-67*).

**As to dependent claim 5,** Chakrabarti teaches training the active probability includes multiplying the active probability indicative of a change in the document by a training probability calculated using a statistical model (*Abstract*).

**As to dependent claim 6,** Chakrabarti teaches training a document probability distribution corresponding to the document address specification to reflect an experience with the document during a plurality of previous crawls, the document probability distribution including a plurality of probabilities; determining from the document probability distribution a probability that the document has changed; and making a determination of whether to access the document in a current crawl based on the probability that the document has changed (*col. 7, lines 3-17 & col. 8, lines 53-67*).

**As to dependent claim 7,** Chakrabarti teaches calculating, based on the experience with the document during a plurality of previous crawls, a discrete random variable distribution that includes a plurality of training probabilities; multiplying each probability in the document probability distribution by a corresponding training probability from the discrete random variable distribution (*col. 7, lines 3-17*).

**As to dependent claim 8,** Chakrabarti teaches the training probabilities are calculated using a Poisson process, the Poisson process including a Poisson equation ( $e^{-r*dt}$ ) and a complementary Poisson equation ( $1-e^{-r*dt}$ ) (*col. 7, lines 3-17*).

**As to dependent claim 9,** Chakrabarti teaches the experience with the document during the plurality of previous crawls is derived from historical information associated with the document address specification (*col. 3, lines 1-30*).

**As to independent claim 10,** Chakrabarti teaches (*col.5, lines 14-60*) a computer-readable medium having computer-executable instructions for retrieving one document in a plurality of documents (*a user can search ... Web pages of interest*) from a remote server (*the Web server*), which when executed comprise:

- maintaining historical information associated with changes to the one document at the remote server (*the Web page table ... indicating the date and time the Web page was last modified by the provider of the content of the page; col.5, lines 29-60*);
- initiating a crawl procedure for retrieving particular documents in the plurality of documents (*collectively downloaded Web pages that are related to a limited number of predefined topics; col.5, line 61-col.6, line 15/ crawl database 30 to retrieve a list of relevant Web pages; col.6, lines 35-51*); and
- determining whether to access the one document from the remote server based on an analysis of the historical information associated with the changes to the one document at the remote server (*the Web page table 32 includes a priority field 42 ... determined in consonance with the topic analyzer 28 to determine whether any changes ... to refresh the entry; col.5, lines 29-60*).

**As to dependent claim 11,** Chakrabarti teaches if the determination to access the one document is positive, identifying the one document for retrieval during the crawl procedure; and attempting to retrieve all documents identified for retrieval during the crawl procedure (*col.11, lines 17-29*).

**As to dependent claim 12,** Chakrabarti teaches computing a probability that the one document has changed since the one document was last retrieved from the remote server (*col. 7, lines 3-17 & col. 8, lines 53-67*).

**As to dependent claim 13,** Chakrabarti teaches beginning with a probability that a pre-defined proportion of documents in the plurality of documents has changed, training the probability that the pre-defined proportion of documents has changed using the historical information associated with the one document to achieve the probability that the one document has changed (*col. 10, lines 18-34 & crawl table entry, Fig. 1*).

**As to dependent claim 14,** Chakrabarti teaches comprising making a random decision to retrieve the one document wherein the random decision is biased by the probability that the one document has changed (*col. 7, lines 18-41*).

**As to dependent claim 15,** Chakrabarti teaches the random decision is further biased by a synchronization level configured to influence the random decision based on a predetermined degree of tolerance for not retrieving the one document if the document is likely to have changed (*col. 7, lines 18-41*).

**As to dependent claim 16,** Chakrabarti teaches the random decision is made by a software routine adapted to simulate a flip of a coin (*col. 7, lines 18-41*).

**As to dependent claim 17,** Chakrabarti teaches the historical information associated with changes to the one document includes a time stamp for the one document, the time stamp being indicative of a last time that the one document was modified when the one document was last retrieved from the remote server; and wherein the analysis includes a comparison of the time

stamp included in the historical information with another time stamp associated with the one document stored on the remote server (*crawl table entry, Fig.1*).

**As to dependent claim 18**, Chakrabarti teaches if the time stamp included in the historical information does not match the other time stamp associated with the one document stored on the remote server, identifying the one document for retrieval during the crawl procedure (*crawl table entry, Fig.1*).

**As to dependent claim 19**, Chakrabarti teaches the historical information associated with changes to the one document includes a hash value associated with the one document, the hash value being a representation of the one document; and wherein the analysis includes a comparison of the hash value included in the historical information with another hash value calculated from information retrieved from the one document stored on the remote server (*crawl table entry, Fig.1*).

**As to dependent claim 20**, Chakrabarti teaches if the hash value included in the historical information does not match the other hash value associated with the one document stored on the remote server, identifying the one document for retrieval during the crawl procedure (*crawl table entry, Fig.1*).

### **Conclusion**

4. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Narendran et al.	U.S Patent No. 6,070,191	issue dated: May 30, 2000
Dillon	U.S Patent No. 6,351,467	issue dated: Feb. 26, 2002
Liddy et al.	U.S Patent No. 6,304,864	issue dated: Oct. 16, 2001
Najork et al.	U.S Patent No. 5,594,660	issue dated: Jul. 17, 2001
Najork et al.	U.S Patent No. 6,351,755	issue dated: Feb. 26, 2002
Sundaresan et al.	U.S Patent No. 6,539,376	issue dated: Mar.25, 2003
Wiener et al.	U.S Patent No. 6,598,051	issue dated: Jul. 22, 2003

5. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Maikhahan Nguyen whose telephone number is (703) 306-0092. The examiner can normally be reached on Monday - Friday from 9:00am – 5:30 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Joseph H. Feild can be reached on (703) 305-9792. The fax phone numbers for the organization where this application or proceeding is assigned are (703) 308-5403 for regular communications and (703) 308-5403 for After Final communications.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 305-9600.

**Contact Information:**

Any response to this action should be mailed to:

Commissioner for Patents  
PO Box 1450  
Alexandria, VA 22313-1450

Or fax to:

AFTER-FINAL faxes must be signed and sent to (703) 746-7238.  
OFFICIAL faxes must be signed and sent to (703) 872-9306.  
NON OFFICIAL faxes should be sent to (703) 746-7240.

Art Unit: 2176

All OFFICIAL faxes will be handled and entered by the docketing personnel. The date of entry will correspond to the actual FAX reception date unless that date is a Saturday, Sunday, or a Federal Holiday within the District of Columbia, in which case the official date of receipt will be the next business day. The application file will be promptly forwarded to the Examiner unless the application file must be sent to another area of the Office, e.g., Finance Division for fee charging, etc.

Hand-delivered responses should be brought to Crystal Park II, 2121 Crystal Drive, Arlington, VA., Sixth Floor (Receptionist). All hand-delivered responses will be handled and entered by the docketing personnel. Please do not hand deliver responses directly to the Examiner.

Maikhahan Nguyen  
September 16, 2003



SANJIV SHAH  
PRIMARY EXAMINER